

The degree distribution and the number of edges between nodes of given degrees in the Buckley–Osthus model of a random web graph

Evgeniy A. Grechnikov*

Abstract

In this paper, we study some important statistics of the random graph $H_{a,k}^{(t)}$ in the Buckley–Osthus model, where t is the number of nodes, kt is the number of edges (so that $k \in \mathbb{N}$), and $a > 0$ is the so-called initial attractiveness of a node. This model is a modification of the well-known Bollobás–Riordan model. First, we find a new asymptotic formula for the expectation of the number $R(d, t)$ of nodes of a given degree d in a graph in this model. Such a formula is known for $a \in \mathbb{N}$ and $d \leq t^{1/100(a+1)}$. Both restrictions are unsatisfactory from theoretical and practical points of view. We completely remove them. Then we calculate the covariances between any two quantities $R(d_1, t)$, $R(d_2, t)$, and using the second moment method we show that $R(d, t)$ is tightly concentrated around its mean for every possible values of d and t . Furthermore, we study a more complicated statistic of the web graph: $X(d_1, d_2, t)$ is the total number of edges between nodes whose degrees are equal to d_1 and d_2 respectively. We also find an asymptotic formula for the expectation of $X(d_1, d_2, t)$ and prove a tight concentration result. Again, we do not impose any substantial restrictions on the values of d_1, d_2 , and t .

1 Introduction

The real world has many interesting structures which can be thought of as graphs. A typical example is the World Wide Web: one can consider web pages to be nodes of a graph and hyperlinks to be edges. One of productive methods for studying these graphs involves investigation of a suitable random graph model.

First models of random graphs were constructed and investigated long ago. Classical models and results are systematized, for example, in [4] and [11]. However, they are not suitable for approximation of dynamically changing and non-uniform networks. In particular, the degree sequences of the graphs in these models are very far from those observed in reality.

Recently other models of random graphs were constructed to more closely match the growth of real networks. One of the first descriptions of such a model belongs to the article [3] by Barabási and Albert. The authors of this article introduced the “preferential attachment” rule. Models following this rule assign the probability of a new edge to a node according to the current degree of this node, so more “popular” nodes are more attractive for new edges.

However, the article [3] did not contain a precise model, leaving some parameters unspecified. Variations of these parameters can significantly change properties of arising graphs, as shown in [5], so one needs something more explicit for theoretical investigations. Bollobás, Riordan et al. proposed an explicit model in [6] based on the preferential attachment rule. In the same article, they rigorously proved a theorem concerning the degree sequence of a graph in this model. Namely, they showed that the number of nodes with degree d in their model decreases proportional to d^{-3} . The same quantity in real networks decreases proportional to $d^{-\gamma}$ with different γ for different networks, following the so called “power law”.

*Research division in Yandex, Moscow.

The Bollobás–Riordan model has only one parameter, a natural number representing the ratio of the number of edges to the number of nodes. Thus, on the one hand, the Bollobás–Riordan model does certainly match some real networks by explaining the power law. But, on the other hand, the number of parameters in this model is small and does not allow to obtain the power law with an exponent, which is not equal to -3 .

In the Bollobás–Riordan model, the probability that a node is target for a new edge is proportional to the degree of this node. In [8] and [9] two groups of researchers independently proposed to add to the model one more parameter — an “initial attractiveness” of a node which is a positive constant not depending on the degree. Equivalently, the probability in the proposed model is a linear function in the degree. However, in the papers [8] and [9], we find only some heuristic arguments.

In [7] Buckley and Osthus gave an explicit construction of the above-described model and rigorously proved a theorem concerning the degree sequence of a graph in this model when all the parameters are natural numbers.

Among many articles in this area, we also quote [12]. The model investigated in this article differs from the Buckley–Osthus model, but the difference is small, so the results are comparable. The article deals with the case when parameters are not necessarily natural. However, the proven theorem only works for fixed degree d when the number of nodes tends to infinity; Bollobás et al. as well as Buckley and Osthus allowed d to grow as some small power of the number of nodes.

There are many other random graph models intended to approximate real networks. We refer the reader to [5] and [10] for surveys of such models and corresponding results.

We study the Buckley–Osthus model of a random graph. Our first goal is to give a significant improvement of the above-mentioned theorem from [7] using a completely different method. We find an asymptotic formula for the expectation of the number of nodes with degree d without any upper bound on d and with an estimation of the error term. We also prove a tight concentration result.

Since the Bollobás–Riordan model is a special case of the Buckley–Osthus model, our results are also applicable to it. So, again, we get a substantial improvement of the main theorem from [6].

Our second goal is to study the following quantity. We fix two numbers d_1 and d_2 . We consider a node with degree d_1 and a node with degree d_2 . Then, we calculate the number of edges between these nodes. When there are several choices for nodes of given degrees, we calculate the mean value. Since the number of nodes with a fixed degree is known to have tight concentration around its expectation, it is sufficient to examine the total number of edges linking a node with degree d_1 and a node with degree d_2 . Here we also obtain an asymptotic formula for the expectation and prove a tight concentration result.

2 The model and formulation of results

The Buckley–Osthus model has two parameters, a natural number k and a positive real number a . The number k is the ratio of the number of edges to the number of nodes. We assume that a and k are constants, so by default all other constants may depend on them. The Bollobás–Riordan model is a special case of this model with $a = 1$.

The model is defined in two stages. At the first stage, a probability space $H_{a,1}^{(t)}$ is constructed. The elements of $H_{a,1}^{(t)}$ are undirected graphs with nodes represented by numbers $1, \dots, t$ and with t edges. The space $H_{a,1}^{(1)}$ contains only one graph with one node and one loop. Any graph in $H_{a,1}^{(t)}$ is obtained from a graph in $H_{a,1}^{(t-1)}$ by adding a new node t and a new edge between t and a node $\gamma \in \{1, \dots, t\}$ so that

$$\Pr(\gamma = s) = \begin{cases} \frac{\deg_{t-1}(s) - 1 + a}{(a+1)t-1}, & 1 \leq s \leq t-1, \\ \frac{a}{(a+1)t-1}, & s = t, \end{cases}$$

where \deg_{t-1} denotes the degree of a node in the graph from $H_{a,1}^{(t-1)}$. At the second stage, a final probability space $H_{a,k}^{(t)}$ is constructed from $H_{a,1}^{(tk)}$ as follows. We take any graph from $H_{a,1}^{(tk)}$. It has kt nodes and kt edges. We identify the nodes $1, \dots, k; k+1, \dots, 2k; \dots$ obtaining t new nodes, and we keep all the edges obtaining multiple edges and even multiple loops.

We study the number of nodes of degree d in $H_{a,k}^{(t)}$ as a function of d and t . We denote this random quantity by $R(d, t)$ and the value of its expectation by $r(d, t) = ER(d, t)$.

If $d < k$, then clearly $R(d, t) = 0$, so it suffices to study the case $d \geq k$. We start by considering $r(d, t)$.

Theorem 1. *Let $d \geq k$. The expected value of $R(d, t)$ is*

$$r(d, t) = \frac{B(d - k + ka, a + 2)}{B(ka, a + 1)}t + O_{a,k}\left(\frac{1}{d}\right).$$

The asymptotic behaviour of the coefficient when d grows is

$$\frac{B(d - k + ka, a + 2)}{B(ka, a + 1)} \sim \frac{\Gamma(a + 2)}{B(ka, a + 1)}d^{-2-a} = (a + 1)\frac{\Gamma(ka + a + 1)}{\Gamma(ka)}d^{-2-a}.$$

A similar result was obtained in [7] (with some factorials instead of Gamma- and Beta-functions). However, for that result, it was essential that $a \in \mathbb{N}$ and $d \leq t^{1/100(a+1)}$. Another result, which can be compared with the one of Theorem 1, is proved in [12]. It concerns a bit different model, but, nevertheless, it is rather close to our investigations. In this result, a can be any positive real (and so analogous Gamma- and Beta-functions appear in its statement). However, its proof essentially uses the assumption that d is just a constant. In our Theorem 1, we do not have any restrictions on d and a , and we use a completely different method to prove it.

In fact, Theorem 1 gives an entire picture of what happens to the quantity $r(d, t)$. If $d = o\left(t^{\frac{1}{a+1}}\right)$, then Theorem 1 yields the main term of $r(d, t)$. If $d = \Omega\left(t^{\frac{1}{a+1}}\right)$, then $r(d, t)$ tends to zero as $t \rightarrow \infty$, which means that with high probability there are no nodes of degree d in a graph in the model.

Now we want to study in detail the quantity $R(d, t)$.

Theorem 2. *Let $d_1 \geq k$ and $d_2 \geq k$. The covariance between $R(d_1, t)$ and $R(d_2, t)$ is*

$$\text{cov}(R(d_1, t), R(d_2, t)) = O_{a,k}\left((d_1^{-2-a} + d_2^{-2-a})t + d_1^{-1}d_2^{-1}\right).$$

Substituting $d_1 = d_2 = d$ in Theorem 2 and using Chebyshev's inequality, we obtain the following result.

Corollary 1. *If $d = d(t) \geq k$ and $\psi(t) \rightarrow \infty$ as $t \rightarrow \infty$, then*

$$\left|R(d, t) - \frac{B(d - k + ka, a + 2)}{B(ka, a + 1)}t\right| \leq \left(\sqrt{d^{-a-2}t} + d^{-1}\right)\psi(t) \quad (1)$$

with probability tending to 1 as $t \rightarrow \infty$.

Let us discuss the meaning of Corollary 1.

When $d \sim Ct^{\frac{1}{a+2}}$ with some constant C , both $r(d, t)$ and $\sqrt{d^{-a-2}t} + d^{-1}$ are $O(1)$. For smaller values of d (i.e., $d = o\left(t^{\frac{1}{a+2}}\right)$), inequality (1) implies the equivalence (with probability tending to 1 as $d, t \rightarrow \infty$)

$$R(d, t) \sim \frac{(a + 1)\Gamma(ka + a + 1)}{\Gamma(ka)}d^{-2-a}t.$$

For larger values of d (i.e., $t^{\frac{1}{a+2}} = o(d)$), inequality (1) means that $R(d, t) = o(1)$. Since $R(d, t)$ is an integer number by definition, $R(d, t) = 0$ (again, with probability tending to 1 as $d, t \rightarrow \infty$). Thus, we have an almost entire picture of what happens to $R(d, t)$.

We also study the total number of edges linking a node with degree d_1 and a node with degree d_2 . We denote this random quantity by $X(d_1, d_2, t)$. When $d_1 = d_2$, we count every edge twice, but do not count loops.

Theorem 3. Let $d_1 \geq k$ and $d_2 \geq k$. There exists a function $c_X(d_1, d_2)$ such that

$$EX(d_1, d_2, t) = c_X(d_1, d_2)t + O_{a,k}(1)$$

and

$$c_X(d_1, d_2) = \frac{\Gamma(d_1 - k + ka)\Gamma(d_2 - k + ka)\Gamma(d_1 + d_2 - 2k + 2ka + 3)}{\Gamma(d_1 - k + ka + 2)\Gamma(d_2 - k + ka + 2)\Gamma(d_1 + d_2 - 2k + 2ka + a + 2)} \times \\ \times ka(a+1) \frac{\Gamma(ka + a + 1)}{\Gamma(ka)} \left(1 + \theta(d_1, d_2) \frac{(d_1 - k + ka + 1)(d_2 - k + ka + 1)}{(d_1 + d_2 - 2k + 2ka + 1)(d_1 + d_2 - 2k + 2ka + 2)} \right),$$

where

$$-4 + \frac{2}{1 + ka} \leq \theta(d_1, d_2) \leq a \frac{\Gamma(ka + 1)\Gamma(2ka + a + 3)}{\Gamma(2ka + 2)\Gamma(ka + a + 2)}.$$

When both d_1 and d_2 grow, the asymptotic behaviour of c_X is

$$c_X(d_1, d_2) = ka(a+1) \frac{\Gamma(ka + a + 1)}{\Gamma(ka)} \frac{(d_1 + d_2)^{1-a}}{d_1^2 d_2^2} \left(1 + O_{a,k} \left(\frac{1}{d_1} + \frac{1}{d_2} + \frac{d_1 d_2}{(d_1 + d_2)^2} \right) \right).$$

Note that the last formula in Theorem 3 does not give an asymptotic behaviour if d_1 and d_2 grow so that $\frac{d_2}{d_1}$ tends to a finite nonzero limit. The precise bounds show that the term $\frac{(d_1 + d_2)^{1-a}}{d_1^2 d_2^2}$ still gives the correct order of growth for c_X , but the coefficient can differ from $ka(a+1) \frac{\Gamma(ka+a+1)}{\Gamma(ka)}$. And in fact, the coefficient differs.

Theorem 4. Let $d_1, d_2 \geq k$, $c > 0$. Then

$$P \left(|X(d_1, d_2, t) - EX(d_1, d_2, t)| \geq c(d_1 + d_2)\sqrt{kt} \right) \leq 2 \exp \left(-\frac{c^2}{8} \right).$$

In particular, if $c(t) \rightarrow \infty$ as $t \rightarrow \infty$, then $|X - EX| < c(t)(d_1 + d_2)\sqrt{kt}$ with probability tending to 1.

From Theorems 3 and 4, we immediately obtain the following assertion.

Corollary 2. If $(d_1 + d_2)^a d_1^2 d_2^2 = o(\sqrt{t})$, then with probability tending to 1 as $d_1, d_2, t \rightarrow \infty$

$$X(d_1, d_2, t) \sim c_X(d_1, d_2)t.$$

The mean value of the number of edges between one node with degree d_1 and another node with degree d_2 is $\frac{X(d_1, d_2, t)}{R(d_1, t)R(d_2, t)}$. Since the quantities $R(d, t)$ and $X(d_1, d_2, t)$ are tightly concentrated around their expectations, the main term of the ratio is

$$\frac{\Gamma(ka + 1)}{(a + 1)\Gamma(ka + a + 1)} \frac{d_1^a d_2^a (d_1 + d_2)^{1-a}}{t}.$$

Again, the constant factor can differ if d_1 and d_2 grow so that $\frac{d_2}{d_1}$ tends to a finite nonzero limit, but the order is correct even in this case.

3 Proof of Theorem 1

For a property P , we denote

$$[P] = \begin{cases} 1, & P \text{ holds,} \\ 0, & \text{otherwise.} \end{cases}$$

First of all, we reformulate the model without references to $H_{a,1}^{(t)}$. The probability space $H_{a,k}^{(1)}$ obviously consists of one graph with 1 node and k loops. The space $H_{a,k}^{(t+1)}$ can be obtained from $H_{a,k}^{(t)}$ by adding to

any graph from $H_{a,k}^{(t)}$ a new node $t + 1$ and k edges in the following k steps. At the i th step, we add one edge between the new node and one of the existing nodes γ . If $\gamma \neq t + 1$, then it corresponds to a group of nodes $\gamma_1, \dots, \gamma_k$ in $H_{a,1}^{(kt+i-1)}$. The sum of degrees of $\gamma_1, \dots, \gamma_k$ equals the degree of γ in the graph before the i th step. We denote this degree by $\deg_{t,i}$. So

$$\Pr(\gamma = s) = \frac{\deg_{t,i}(s) + k(a-1)}{(a+1)(kt+i)-1}, \quad 1 \leq s \leq t.$$

If $\gamma = t + 1$, the corresponding group in $H_{a,1}^{(kt+i-1)}$ has only $i - 1$ nodes. Hence,

$$\Pr(\gamma = t + 1) = \frac{\deg_{t,i}(t + 1) + (i - 1)(a - 1) + a}{(a + 1)(kt + i) - 1}.$$

We want to express any value $r(d, t)$ in terms of some values with smaller t . Let us consider the transition from $H_{a,k}^{(t)}$ to $H_{a,k}^{(t+1)}$. Let $r(d, t, i)$ denote the average number of nodes of degree d , not including the last node $t + 1$, before the i th step, and $r(d, t, i + 1)$ — the similar number after the i th step. Let γ be a head of the edge added in the i th step. Then,

$$\begin{aligned} r(d, t, i + 1) &= \sum_{s=1}^t \Pr(\deg_{t,i+1}(s) = d) = \sum_{s=1}^t (\Pr(\deg_{t,i+1}(s) = d, \gamma = s) + \\ &\quad + \Pr(\deg_{t,i+1}(s) = d, \gamma \neq s)) = \sum_{s=1}^t (\Pr(\deg_{t,i}(s) = d - 1, \gamma = s) + \\ &\quad + \Pr(\deg_{t,i}(s) = d, \gamma \neq s)) = \sum_{s=1}^t \left(\Pr(\deg_{t,i}(s) = d - 1) \frac{(d - 1) + k(a - 1)}{(a + 1)(kt + i) - 1} + \right. \\ &\quad \left. + \Pr(\deg_{t,i}(s) = d) \left(1 - \frac{d + k(a - 1)}{(a + 1)(kt + i) - 1} \right) \right) = \\ &= r(d - 1, t, i) \frac{(d - 1) + k(a - 1)}{(a + 1)(kt + i) - 1} + r(d, t, i) \left(1 - \frac{d + k(a - 1)}{(a + 1)(kt + i) - 1} \right). \end{aligned} \quad (2)$$

By definition,

$$r(d, t) = r(d, t, 1), \quad r(d, t + 1) = r(d, t, k + 1) + \Pr(\deg_{t,k+1}(t + 1) = d). \quad (3)$$

The function $r(d, t)$ is completely determined by the equations (2), (3) and the starting condition

$$r(d, 1) = [d = 2k]. \quad (4)$$

The equation (3) includes the function $\Pr(\deg_{t,k+1}(t + 1) = d)$. Obviously,

$$\Pr(\deg_{t,k+1}(t + 1) = d) = 0, \quad d < k \text{ or } d > 2k. \quad (5)$$

The minimal value $\deg_{t,k+1}(t + 1) = k$ is obtained when no one of the k edges is a loop. In this case, $\deg_{t,i}(t + 1) = i - 1$ for all i , so

$$\Pr(\deg_{t,k+1}(t + 1) = k) = \prod_{i=1}^k \left(1 - \frac{ia}{(a + 1)(kt + i) - 1} \right) = 1 + O\left(\frac{1}{t}\right).$$

(Note that a constant in $O()$ depends on a and k).

Because $\sum_{d=k}^{2k} \Pr(\deg_{t,k+1} = d) = 1$ and $\Pr(\deg_{t,k+1} = d) \geq 0$, we get

$$\Pr(\deg_{t,k+1}(t + 1) = d) = O\left(\frac{1}{t}\right), \quad k < d \leq 2k.$$

Since d is bounded in the last equality, its right hand side can be equivalently written as $O\left(\frac{1}{d^2 t}\right)$.

Let

$$c(d) = \begin{cases} \frac{B(d-k+ka, a+2)}{B(ka, a+1)}, & d \geq k, \\ 0, & d < k. \end{cases} \quad (6)$$

If $d > k$, then

$$\begin{aligned} \frac{c(d-1)}{c(d)} &= \frac{B(d-1-k+ka, a+2)}{B(d-k+ka, a+2)} = \\ &= \frac{\Gamma(d-1-k+ka)/\Gamma(d+1-k+ka+a)}{\Gamma(d-k+ka)/\Gamma(d-k+ka+a+2)} = \frac{d+1-k+ka+a}{d-1-k+ka}. \end{aligned}$$

Also

$$c(k) = \frac{B(ka, a+2)}{B(ka, a+1)} = \frac{\Gamma(a+2)/\Gamma(ka+a+2)}{\Gamma(a+1)/\Gamma(ka+a+1)} = \frac{a+1}{ka+a+1}.$$

In particular, $c(d-1) > c(d)$, so $c(d) < c(k) < 1$ for all $d \geq k$.

For the rest of the proof, we will assume that $d \geq k$. Note that of course this does not imply $d-1 \geq k$.

When d grows, the asymptotic behaviour of $c(d)$ is

$$\begin{aligned} \ln c(d) &= \ln \frac{\Gamma(a+2)}{B(ka, a+1)} \frac{\Gamma(d-k+ka)}{\Gamma(d-k+ka+a+2)} = \ln \frac{\Gamma(a+2)}{B(ka, a+1)} + \\ &+ (d-k+ka)(\ln(d-k+ka) - 1) - (d-k+ka+a+2)(\ln(d-k+ka+a+2) - 1) + O\left(\frac{1}{d}\right) = \\ &= \ln \frac{\Gamma(a+2)}{B(ka, a+1)} + (d-k+ka)(\ln d + \frac{-k+ka}{d} - 1) - (d-k+ka+a+2)(\ln d + \frac{-k+ka+a+2}{d} - 1) \\ &\quad + O\left(\frac{1}{d}\right) = \ln \frac{\Gamma(a+2)}{B(ka, a+1)} - (a+2) \ln d + O\left(\frac{1}{d}\right), \\ c(d) &= \frac{\Gamma(a+2)}{B(ka, a+1)} d^{-2-a} \left(1 + O\left(\frac{1}{d}\right)\right). \end{aligned} \quad (7)$$

Let

$$\tilde{r}(d, t, i) = r(d, t, i) - c(d) \left(t + \frac{i}{k} - \frac{1}{k(a+1)}\right).$$

It is easy to see that the theorem is equivalent to $\tilde{r}(d, t, i) = O(1)$. Using (2), we obtain

$$\begin{aligned} \tilde{r}(d-1, t, i) &\frac{(d-1)+k(a-1)}{(a+1)(kt+i)-1} + \tilde{r}(d, t, i) \left(1 - \frac{d+k(a-1)}{(a+1)(kt+i)-1}\right) = \\ &= r(d, t, i+1) - \left(t + \frac{i}{k} - \frac{1}{k(a+1)}\right) \left(c(d-1) \frac{(d-1)+k(a-1)}{(a+1)(kt+i)-1} + \right. \\ &+ c(d) \left(1 - \frac{d+k(a-1)}{(a+1)(kt+i)-1}\right) \Big) = r(d, t, i+1) - \left(t + \frac{i}{k} - \frac{1}{k(a+1)}\right) c(d) \times \\ &\times \frac{(1-[d=k])(d+1-k+ka+a) + (a+1)(kt+i) - 1 - (d+k(a-1))}{(a+1)(kt+i)-1} = \\ &= r(d, t, i+1) - c(d) \frac{(a+1)(kt+i+1) - 1 - [d=k](1+ka+a)}{k(a+1)} = \\ &= \tilde{r}(d, t, i+1) + [d=k]c(k) \frac{1+ka+a}{k(a+1)} = \tilde{r}(d, t, i+1) + \frac{[d=k]}{k}. \end{aligned} \quad (8)$$

Let $C = C(a, k)$ be a sufficiently large constant which will be determined later. We claim that

$$\left| \tilde{r}(d, t, i) + (i-1) \frac{[d=k]}{k} \right| \leq \frac{C}{d+ka} \left(1 - \frac{\min\{1, ka\}}{(a+2)(t+1)(d+ka)}\right)^{i-1} \quad (9)$$

for all $i = 1, \dots, k+1$ and for all natural $d \geq k$ and t . Note that this implies $\tilde{r}(d, t, i) = O\left(\frac{1}{d}\right)$ and Theorem 1.

The equations (2), (3), (4), (5) imply that $r(d, t, i) = 0$ if $d > kt + i - 1 + k$. In this case, using (7), we obtain

$$\tilde{r}(d, t, i) = -c(d) \left(t + \frac{i}{k} - \frac{1}{k(a+1)} \right) = O(d^{-2-a}t) = O(d^{-1-a}),$$

so if $d > kt + i - 1 + k$, (9) is true for all sufficiently large values of C .

Now assume $d \leq kt + i - 1 + k$. We will prove (9) by induction on t and, for fixed t , on i . The basis of induction $t = 1, \dots, 1 + \lfloor \frac{1}{ka} \rfloor$ and any $i = 1, \dots, k+1$ obviously holds for all sufficiently large values of C .

Now let $t \geq 2 + \lfloor \frac{1}{ka} \rfloor$ and let (9) hold for $t-1$. Using (3), we obtain

$$\tilde{r}(d, t, 1) = \tilde{r}(d, t-1, k+1) + \Pr(\deg_{t-1, k+1}(t) = d) = \tilde{r}(d, t-1, k+1) + [d = k] + O\left(\frac{1}{d^2t}\right).$$

Therefore,

$$\begin{aligned} |\tilde{r}(d, t, 1)| &\leq \frac{C}{d+ka} \left(1 - \frac{\min\{1, ka\}}{(a+2)(t+1)(d+ka)} \right)^k + O\left(\frac{1}{d^2t}\right) \leq \\ &\leq \frac{C}{d+ka} - \frac{C \min\{1, ka\}/(a+2)}{(t+1)(d+ka)^2} + O\left(\frac{1}{d^2t}\right). \end{aligned}$$

Thus, induction step on t is proved for all sufficiently large values of C .

Finally, let $t \geq 2 + \lfloor \frac{1}{ka} \rfloor > 1 + \frac{1}{ka}$, $i > 1$ and let (9) hold for $i-1$. We temporarily denote $T = (a+1)(kt+i-1)-1$. Note that T depends on t and i , but not on d . From (8) we obtain

$$\begin{aligned} \tilde{r}(d, t, i) + (i-1)\frac{[d=k]}{k} &= (i-2)\frac{[d=k]}{k} + \\ &+ \tilde{r}(d-1, t, i-1)\frac{(d-1)+k(a-1)}{T} + \tilde{r}(d, t, i-1) \left(1 - \frac{d+k(a-1)}{T} \right) = \\ &= \left(\tilde{r}(d-1, t, i-1) + (i-2)\frac{[d-1=k]}{k} \right) \frac{(d-1)+k(a-1)}{T} + \\ &+ \left(\tilde{r}(d, t, i-1) + (i-2)\frac{[d=k]}{k} \right) \left(1 - \frac{d+k(a-1)}{T} \right) + [k \leq d \leq k+1]O\left(\frac{1}{t}\right). \end{aligned}$$

The remainder term $[k \leq d \leq k+1]O\left(\frac{1}{t}\right)$ can be written as $O\left(\frac{1}{d^2t}\right)$. The assumptions $d \leq kt + i - 1 + k$ and $t > 1 + \frac{1}{ka}$ imply that $1 - \frac{d+k(a-1)}{(a+1)(kt+i-1)-1} \geq 0$. If $d > k$, then

$$\begin{aligned} \frac{C}{d-1+ka} \frac{(d-1)+k(a-1)}{T} + \frac{C}{d+ka} \left(1 - \frac{d+k(a-1)}{T} \right) &= \\ &= \frac{C}{d+ka} \left(1 - \frac{1}{T} \left(-\frac{d+ka}{d-1+ka} ((d-1)+k(a-1)) + d+k(a-1) \right) \right) = \\ &= \frac{C}{d+ka} \left(1 - \frac{1}{T} \left(1 - \frac{(d-1)+k(a-1)}{d-1+ka} \right) \right) = \frac{C}{d+ka} \left(1 - \frac{k}{T(d-1+ka)} \right) \leq \\ &\leq \frac{C}{d+ka} \left(1 - \frac{k \min\{1, ka\}}{T(d+ka)} \right). \end{aligned}$$

If $d = k$, then

$$\frac{C}{d+ka} \left(1 - \frac{d+k(a-1)}{T} \right) = \frac{C}{d+ka} \left(1 - \frac{ka(k+ka)}{T(d+ka)} \right) \leq \frac{C}{d+ka} \left(1 - \frac{k \min\{1, ka\}}{T(d+ka)} \right).$$

In both cases,

$$\left| \tilde{r}(d, t, i) + (i-1) \frac{[d=k]}{k} \right| \leq \frac{C}{d+ka} \left(1 - \frac{\min\{1, ka\}}{(a+2)(t+1)(d+ka)} \right)^{i-2} \left(1 - \frac{k \min\{1, ka\}}{T(d+ka)} \right) + O\left(\frac{1}{d^2 t}\right). \quad (10)$$

Note that

$$1 - \frac{k \min\{1, ka\}}{((a+1)(kt+i-1)-1)(d+ka)} = \left(1 - \frac{k \min\{1, ka\}}{(a+2)(kt+k)(d+ka)} \right) - \frac{k \min\{1, ka\}}{(a+1)(a+2)kt(d+ka)} \left(1 + O\left(\frac{1}{t}\right) \right),$$

where the left part is strictly less than the first term in the right part, so that the second term in the right part is positive. Now (10) implies (9) for all sufficiently large values of C .

Theorem 1 is proved.

4 Proof of Theorem 2

By definition and linearity of expectation,

$$\begin{aligned} \text{cov}(R(d_1, t), R(d_2, t)) &= E(R(d_1, t)R(d_2, t)) - r(d_1, t)r(d_2, t) = \\ &= E \sum_{\substack{s_1, s_2=1 \\ s_1 \neq s_2}}^t [\deg s_1 = d_1, \deg s_2 = d_2] - r(d_1, t)r(d_2, t) = \\ &= \sum_{s_1 \neq s_2} \Pr(\deg s_1 = d_1, \deg s_2 = d_2) + [d_1 = d_2]r(d_1, t) - r(d_1, t)r(d_2, t). \end{aligned} \quad (11)$$

We will estimate the sum

$$r_2(d_1, d_2, t) = \sum_{\substack{s_1, s_2=1 \\ s_1 \neq s_2}}^t \Pr(\deg s_1 = d_1, \deg s_2 = d_2)$$

as we have done it for the function $r(d, t)$ in the proof of Theorem 1.

As with $r(d, t)$, we define a function $r_2(d_1, d_2, t, i)$ as the value of $r_2(d_1, d_2, t)$ before the i th step in the transition from $H_{a,k}^{(t)}$ to $H_{a,k}^{(t+1)}$. The recurrent equation is deduced similarly to (2). For fixed s_1 and s_2 , there are three non-intersecting cases: $\gamma = s_1$, $\gamma = s_2$, and $\gamma \notin \{s_1, s_2\}$. In the first case, we get

$$\begin{aligned} \Pr(\deg_{t,i+1}(s_1) = d_1, \deg_{t,i+1}(s_2) = d_2, \gamma = s_1) &= \\ &= \Pr(\deg_{t,i}(s_1) = d_1 - 1, \deg_{t,i+1}(s_2) = d_2, \gamma = s_1) = \\ &= \Pr(\deg_{t,i}(s_1) = d_1 - 1, \deg_{t,i+1}(s_2) = d_2) \frac{d_1 - 1 + k(a-1)}{(a+1)(kt+i) - 1}. \end{aligned}$$

The second case is the same with d_1 and d_2 interchanged. In the third case, we get

$$\begin{aligned} \Pr(\deg_{t,i+1}(s_1) = d_1, \deg_{t,i+1}(s_2) = d_2, \gamma \neq s_1, \gamma \neq s_2) &= \\ &= \Pr(\deg_{t,i}(s_1) = d_1, \deg_{t,i+1}(s_2) = d_2, \gamma \neq s_1, \gamma \neq s_2) = \\ &= \Pr(\deg_{t,i}(s_1) = d_1, \deg_{t,i+1}(s_2) = d_2) \left(1 - \frac{d_1 + k(a-1) + d_2 + k(a-1)}{(a+1)(kt+i) - 1} \right), \end{aligned}$$

so the final formula is

$$\begin{aligned}
r_2(d_1, d_2, t, i+1) &= r_2(d_1-1, d_2, t, i) \frac{(d_1-1) + k(a-1)}{(a+1)(kt+i)-1} + \\
&\quad + r_2(d_1, d_2-1, t, i) \frac{(d_2-1) + k(a-1)}{(a+1)(kt+i)-1} + \\
&\quad + r_2(d_1, d_2, t, i) \left(1 - \frac{d_1 + d_2 + 2k(a-1)}{(a+1)(kt+i)-1}\right). \quad (12)
\end{aligned}$$

By definition,

$$r_2(d_1, d_2, t) = r_2(d_1, d_2, t, 1),$$

$$\begin{aligned}
r_2(d_1, d_2, t+1) &= r_2(d_1, d_2, t, k+1) + \\
&\quad + \sum_{s=1}^t \Pr(\deg_{t,k+1}(s) = d_1, \deg_{t,k+1}(t+1) = d_2) + \\
&\quad + \sum_{s=1}^t \Pr(\deg_{t,k+1}(s) = d_2, \deg_{t,k+1}(t+1) = d_1), \quad (13)
\end{aligned}$$

and the starting condition is

$$r_2(d_1, d_2, 1) = 0.$$

The equation (12) includes a function

$$r'_2(d_1, d_2, t, i) = \sum_{s=1}^t \Pr(\deg_{t,i}(s) = d_1, \deg_{t,i}(t) = d_2) \quad (14)$$

(and the same function with swapped arguments), so we will first estimate r'_2 . Again, we write a recurrent equation. For fixed s , there are three non-intersecting cases: $\gamma = s$, $\gamma = t+1$, $\gamma \notin \{s, t+1\}$. If $\deg_{t,i+1}(s) = d_1$, then $\deg_{t,i}(s)$ equals $d_1 - 1$ in the first case and d_1 in the two other cases. If $\deg_{t,i+1}(t+1) = d_2$, then $\deg_{t,i}(t+1) = d_2 - 2$ in the second case and $d_2 - 1$ in the two other cases. Calculating probabilities, we obtain

$$\begin{aligned}
r'_2(d_1, d_2, t, i+1) &= r'_2(d_1-1, d_2-1, t, i) \frac{(d_1-1) + k(a-1)}{(a+1)(kt+i)-1} + \\
&\quad + r'_2(d_1, d_2-2, t, i) \frac{(d_2-2) + (i-1)(a-1) + a}{(a+1)(kt+i)-1} + \\
&\quad + r'_2(d_1, d_2-1, t, i) \left(1 - \frac{d_1 + k(a-1) + (d_2-1) + (i-1)(a-1) + a}{(a+1)(kt+i)-1}\right).
\end{aligned}$$

Before the 1st step, the node t has degree 0, so

$$r'_2(d_1, d_2, t, 1) = \sum_{s=1}^t \Pr(\deg_{t,1}(s) = d_1) [d_2 = 0] = [d_2 = 0] r(d_1, t).$$

We continue to use notation (6). Obviously, $r'_2(d_1, d_2, t, i) = 0$ when $d_2 > 2(i-1)$ or $d_1 \geq 2(kt+i)$. If $d_1 < 2(kt+i)$ and $d_2 \leq 2(i-1)$, then $\frac{d_1+d_2}{t} \cdot O(d_1^{-a-2}t) = O(d_1^{-a-1}) = O(d_1^{-1})$ and $\frac{d_1+d_2}{t} \cdot O(d_1^{-1}) = O(d_1^{-1})$. Now it is easy to see that

$$r'_2(d_1, d_2, t, i) = [d_2 = i-1] c(d_1) t + O(d_1^{-1}). \quad (15)$$

Since $r'_2(d_1, d_2, t, i) = 0$ when $d_2 \geq 2k$, the remainder term in (15) is zero when $d_2 \geq 2k$ and can be written as $O\left(\frac{1}{d_1 d_2^{a+1}}\right)$.

Since $r(d_1, t) = 0$ when $d_1 > kt + k$, $r'_2(d_1, d_2, t, i) = 0$ when $d_1 + d_2 > kt + k + 2(i - 1)$. In particular, $r'_2(d_1, d_2, t, k + 1) + r'_2(d_2, d_1, t, k + 1) = 0$ when $d_1 + d_2 > kt + 2k$. By (12) and (13), $r_2(d_1, d_2, t, i) = 0$ when $d_1 + d_2 > kt + 2k + (i - 1)$.

For the rest of the proof, we will assume that $d_1 \geq k$ and $d_2 \geq k$. Note that of course this does not imply $d_1 - 1 \geq k$ and $d_2 - 1 \geq k$.

Let

$$\tilde{r}_2(d_1, d_2, t, i) = r_2(d_1, d_2, t, i) - c(d_1)c(d_2) \left(t + \frac{i}{k} - \frac{1}{k(a+1)} \right) \left(t + \frac{i+1}{k} - \frac{1}{k(a+1)} \right).$$

We temporarily denote $T = kt + i - \frac{1}{a+1}$. We express r_2 in terms of \tilde{r}_2 and use (12). In the expression, there are terms with \tilde{r}_2 with various arguments. Now we transform the terms without \tilde{r}_2 from the right part of (12).

$$\begin{aligned} & \frac{T}{k} \frac{T+1}{k} \left(c(d_1 - 1)c(d_2) \frac{(d_1 - 1) + k(a - 1)}{(a + 1)T} + \right. \\ & \quad \left. + c(d_1)c(d_2 - 1) \frac{(d_2 - 1) + k(a - 1)}{(a + 1)T} + c(d_1)c(d_2) \left(1 - \frac{d_1 + d_2 + 2k(a - 1)}{(a + 1)T} \right) \right) = \\ & = \frac{T}{k} \frac{T+1}{k} c(d_1)c(d_2) \left((1 - [d_1 = k]) \frac{d_1 + 1 - k + ka + a}{(a + 1)T} + \right. \\ & \quad \left. + (1 - [d_2 = k]) \frac{d_2 + 1 - k + ka + a}{(a + 1)T} + 1 - \frac{d_1 + d_2 + 2k(a - 1)}{(a + 1)T} \right) = \\ & = \frac{T}{k} \frac{T+1}{k} c(d_1)c(d_2) \left(\frac{2(a + 1) + (a + 1)T - [d_1 = k](1 + ka + a) - [d_2 = k](1 + ka + a)}{(a + 1)T} \right) = \\ & = \frac{T+1}{k} \frac{T+2}{k} c(d_1)c(d_2) - [d_1 = k] \frac{(T+1)c(d_2)}{k^2} - [d_2 = k] \frac{(T+1)c(d_1)}{k^2}. \end{aligned}$$

The first term equals the term without \tilde{r}_2 in the left part, so

$$\begin{aligned} \tilde{r}_2(d_1, d_2, t, i + 1) &= \tilde{r}_2(d_1 - 1, d_2, t, i) \frac{(d_1 - 1) + k(a - 1)}{(a + 1)(kt + i) - 1} + \\ &+ \tilde{r}_2(d_1, d_2 - 1, t, i) \frac{(d_2 - 1) + k(a - 1)}{(a + 1)(kt + i) - 1} + \tilde{r}_2(d_1, d_2, t, i) \left(1 - \frac{d_1 + d_2 + 2k(a - 1)}{(a + 1)(kt + i) - 1} \right) - \\ &- [d_1 = k] \frac{(kt + i - \frac{1}{a+1} + 1) c(d_2)}{k^2} - [d_2 = k] \frac{(kt + i - \frac{1}{a+1} + 1) c(d_1)}{k^2}. \end{aligned}$$

Relations (13) and (15) imply

$$\tilde{r}_2(d_1, d_2, t, 1) = \tilde{r}_2(d_1, d_2, t - 1, k + 1) + [d_2 = k]c(d_1)t + [d_1 = k]c(d_2)t + O\left(\frac{1}{d_1^{a+1}d_2} + \frac{1}{d_1d_2^{a+1}}\right).$$

Let

$$c_1(d'_1, d'_2) = \begin{cases} \frac{\Gamma(d'_1 - k + ka)}{(d'_2 + k(a - 1))\Gamma(d'_1 - k + ka + a + 1)}, & d'_1 \geq k, d'_2 \geq k, \\ 0, & d'_1 < k \text{ or } d'_2 < k. \end{cases}$$

By definition, for $d_1 > k$,

$$\frac{c_1(d_1 - 1, d_2)}{c_1(d_1, d_2)} = \frac{d_1 - k + ka + a}{d_1 - k + ka - 1}.$$

For $d_2 > k$,

$$\frac{c_1(d_1, d_2 - 1)}{c_1(d_1, d_2)} = \frac{d_2 + k(a - 1)}{d_2 - 1 + k(a - 1)}.$$

Similar to (7),

$$c_1(d_1, d_2) = \frac{d_1^{-1-a}}{d_2 + k(a-1)} \left(1 + O\left(\frac{1}{d_1}\right) \right).$$

Moreover,

$$\begin{aligned} \frac{c_1(d_1, d_2)}{c(d_1)} &= \frac{B(ka, a+1)}{\Gamma(a+2)} \frac{\Gamma(d_1 - k + ka)\Gamma(d_1 - k + ka + a + 2)}{(d_2 + k(a-1))\Gamma(d_1 - k + ka + a + 1)\Gamma(d_1 - k + ka)} = \\ &= \frac{B(ka, a+1)}{\Gamma(a+2)} \frac{d_1 - k + ka + a + 1}{d_2 + k(a-1)}. \end{aligned}$$

Let $C = C(a, k)$ be a sufficiently large constant which will be determined later. We claim that

$$\begin{aligned} \left| \tilde{r}_2(d_1, d_2, t, i) + \frac{i-1}{k}([d_2 = k]c(d_1) + [d_1 = k]c(d_2))t \right| &\leq \\ &\leq C(c_1(d_1, d_2) + c_1(d_2, d_1)) \left(kt + \frac{a + \frac{1}{2}}{a+1}i \right) \quad (16) \end{aligned}$$

for all $i = 1, \dots, k+1$ and for all natural $d_1 \geq k, d_2 \geq k, t$. Since both parts of (16) are symmetric in d_1 and d_2 , it is sufficient to consider the case $d_1 \leq d_2$.

If $d_1 + d_2 > kt + 2k + (i-1)$, then $r_2(d_1, d_2, t, i) = 0$ and

$$\begin{aligned} \frac{|\tilde{r}_2(d_1, d_2, t, i)|}{c_1(d_1, d_2)t} &= \frac{c(d_1)c(d_2) \left(t + \frac{i}{k} - \frac{1}{k(a+1)} \right) \left(t + \frac{i+1}{k} - \frac{1}{k(a+1)} \right)}{c_1(d_1, d_2)t} = \\ &= O\left(\frac{d_2 + k(a-1)}{d_1 - k + ka + a + 1} c(d_2)t \right) = O\left(\frac{t}{d_1 d_2^{1+a}} \right). \end{aligned}$$

Since $d_2 \geq \frac{d_1+d_2}{2} \geq \frac{k}{2}t + k$, the right part is bounded. Obviously, $[d_2 = k] = 0$ and $[d_1 = k]c(d_2)t = O\left(\frac{t}{d_2^{2+a}}\right)$ is bounded too. Thus (16) holds when $d_1 + d_2 > kt + 2k + (i-1)$ for all sufficiently large values of C .

Let $d_1 + d_2 \leq kt + 2k + (i-1)$. We will prove (16) by induction on t and, for fixed t , on i . The basis of induction $t = 1, \dots, 2 + \lfloor \frac{1}{ka} \rfloor$ and any $i = 1, \dots, k+1$ obviously holds for all sufficiently large values of C .

Let $t \geq 3 + \lfloor \frac{1}{ka} \rfloor$ and let (16) hold for $t-1$. We continue to use the restriction $d_1 \leq d_2$. Thus,

$$\begin{aligned} |\tilde{r}_2(d_1, d_2, t, 1)| &= \left| \tilde{r}_2(d_1, d_2, t-1, k+1) + [d_2 = k]c(d_1)t + [d_1 = k]c(d_2)t + O\left(\frac{1}{d_1^{a+1}d_2}\right) \right| \leq \\ &\leq C(c_1(d_1, d_2) + c_1(d_2, d_1)) \left(k(t-1) + \frac{a + \frac{1}{2}}{a+1}(k+1) \right) + O\left(\frac{1}{d_1^{a+1}d_2}\right). \end{aligned}$$

Since $c_1(d_1, d_2) = O\left(\frac{1}{d_1^{a+1}d_2}\right)$, the right part is less than $C(c_1(d_1, d_2) + c_1(d_2, d_1)) \left(kt + \frac{a + \frac{1}{2}}{a+1} \right)$ for all sufficiently large values of C . This completes the induction on t .

Let $t \geq 3 + \lfloor \frac{1}{ka} \rfloor > 2 + \frac{1}{ka}$, $i > 1$ and let (16) hold for $i-1$. Then,

$$\begin{aligned} &\tilde{r}_2(d_1, d_2, t, i) + \frac{i-1}{k}([d_2 = k]c(d_1) + [d_1 = k]c(d_2))t = \\ &= \tilde{r}_2(d_1 - 1, d_2, t, i-1) \frac{(d_1 - 1) + k(a-1)}{(a+1)(kt + i - 1) - 1} + \tilde{r}_2(d_1, d_2 - 1, t, i-1) \frac{(d_2 - 1) + k(a-1)}{(a+1)(kt + i - 1) - 1} + \\ &+ \tilde{r}_2(d_1, d_2, t, i-1) \left(1 - \frac{d_1 + d_2 + 2k(a-1)}{(a+1)(kt + i - 1) - 1} \right) - [d_1 = k] \frac{(kt + i - 1 - \frac{1}{a+1} + 1) c(d_2)}{k^2} - \\ &- [d_2 = k] \frac{(kt + i - 1 - \frac{1}{a+1} + 1) c(d_1)}{k^2} + \frac{i-1}{k}([d_2 = k]c(d_1) + [d_1 = k]c(d_2))t = \end{aligned}$$

$$\begin{aligned}
&= \left(\tilde{r}_2(d_1-1, d_2, t, i-1) + \frac{i-2}{k}([d_2=k]c(d_1-1) + [d_1-1=k]c(d_2))t \right) \frac{(d_1-1) + k(a-1)}{(a+1)(kt+i-1)-1} + \\
&+ \left(\tilde{r}_2(d_1, d_2-1, t, i-1) + \frac{i-2}{k}([d_2-1=k]c(d_1) + [d_1=k]c(d_2-1))t \right) \frac{(d_2-1) + k(a-1)}{(a+1)(kt+i-1)-1} + \\
&+ \left(\tilde{r}_2(d_1, d_2, t, i-1) + \frac{i-2}{k}([d_2=k]c(d_1) + [d_1=k]c(d_2))t \right) \left(1 - \frac{d_1+d_2+2k(a-1)}{(a+1)(kt+i-1)-1} \right) + \\
&\quad + [d_1 \leq k+1]O(c(d_2)) + [d_2 \leq k+1]O(c(d_1)).
\end{aligned}$$

The assumptions $d_1 + d_2 \leq kt + 2k + (i-1)$ and $t > 2 + \frac{1}{ka}$ imply that $1 - \frac{d_1+d_2+2k(a-1)}{(a+1)(kt+i-1)-1} \geq 0$. By the induction hypothesis

$$\begin{aligned}
&\left| \tilde{r}_2(d_1, d_2, t, i) + \frac{i-1}{k}([d_2=k]c(d_1) + [d_1=k]c(d_2))t \right| \leq \\
&\leq C \left(kt + \frac{a+\frac{1}{2}}{a+1}(i-1) \right) \left((c_1(d_1-1, d_2) + c_1(d_2, d_1-1)) \frac{(d_1-1) + k(a-1)}{(a+1)(kt+i-1)-1} + \right. \\
&\quad + (c_1(d_1, d_2-1) + c_1(d_2-1, d_1)) \frac{(d_2-1) + k(a-1)}{(a+1)(kt+i-1)-1} + \\
&\quad \left. + (c_1(d_1, d_2) + c_1(d_2, d_1)) \left(1 - \frac{d_1+d_2+2k(a-1)}{(a+1)(kt+i-1)-1} \right) \right) + \\
&\quad + [d_1 \leq k+1]O(c(d_2)) + [d_2 \leq k+1]O(c(d_1)).
\end{aligned}$$

Since

$$\begin{aligned}
&\frac{c_1(d_1-1, d_2)}{c_1(d_1, d_2)} \frac{(d_1-1) + k(a-1)}{(a+1)(kt+i-1)-1} + \frac{c_1(d_1, d_2-1)}{c_1(d_1, d_2)} \frac{(d_2-1) + k(a-1)}{(a+1)(kt+i-1)-1} + \\
&\quad + 1 - \frac{d_1+d_2+2k(a-1)}{(a+1)(kt+i-1)-1} = (1 - [d_1=k]) \frac{d_1-k+ka+a}{(a+1)(kt+i-1)-1} + \\
&\quad + (1 - [d_2=k]) \frac{d_2+k(a-1)}{(a+1)(kt+i-1)-1} + 1 - \frac{d_1+d_2+2k(a-1)}{(a+1)(kt+i-1)-1} \leq 1 + \frac{a}{(a+1)(kt+i-1)-1},
\end{aligned}$$

we have

$$\begin{aligned}
&\left| \tilde{r}_2(d_1, d_2, t, i) + \frac{i-1}{k}([d_2=k]c(d_1) + [d_1=k]c(d_2))t \right| \leq \\
&\leq C(c_1(d_1, d_2) + c_1(d_2, d_1)) \left(kt + \frac{a+\frac{1}{2}}{a+1}(i-1) \right) \left(1 + \frac{a}{(a+1)(kt+i-1)-1} \right) + \\
&\quad + [d_1 \leq k+1]O(c(d_2)) + [d_2 \leq k+1]O(c(d_1)). \quad (17)
\end{aligned}$$

Since

$$\begin{aligned}
&\left(kt + \frac{a+\frac{1}{2}}{a+1}i \right) - \left(kt + \frac{a+\frac{1}{2}}{a+1}(i-1) \right) \left(1 + \frac{a}{(a+1)(kt+i-1)-1} \right) = \frac{a+\frac{1}{2}}{a+1} - \\
&\quad - \left(kt + \frac{a+\frac{1}{2}}{a+1}(i-1) \right) \frac{a}{(a+1)(kt+i-1)-1} = \frac{1}{(a+1)(kt+i-1)-1} \left(\left(a + \frac{1}{2} \right) (kt+i-1) - \right. \\
&\quad \left. - \frac{a+\frac{1}{2}}{a+1} - a kt - a \frac{a+\frac{1}{2}}{a+1}(i-1) \right) = \frac{1}{(a+1)(kt+i-1)-1} \left(\frac{kt}{2} + \frac{a+\frac{1}{2}}{a+1}(i-2) \right)
\end{aligned}$$

is always positive and tends to a nonzero constant limit as t grows, it is bounded from below by a positive constant. Therefore, for all sufficiently large values of C , the inequality (17) implies the inductive step by i , and so (16) holds.

As a consequence of (16), we obtain

$$\tilde{r}_2(d_1, d_2, t, i) = O\left(\frac{t}{d_1^{a+1}d_2} + \frac{t}{d_1d_2^{a+1}}\right).$$

The proven bound, the representation (11) and Theorem 1 give the following bound:

$$\text{cov}(R(d_1, t), R(d_2, t)) = O\left(\frac{t}{d_1^{a+1}d_2} + \frac{t}{d_1d_2^{a+1}}\right) + O(d_1^{-2-a}t) + O(d_2^{-2-a}t) + O\left(\frac{1}{d_1d_2}\right).$$

If $d_1 \leq d_2$, the maximum among the first three terms on the right-hand side is $O(d_1^{-2-a}t)$; otherwise, the maximum is $O(d_2^{-2-a}t)$. This proves Theorem 2.

5 Proof of Theorem 3

We will use the notation $N(s_1, s_2)$ for the number of edges between nodes s_1 and s_2 . As usual, $N_{t,i}(s_1, s_2)$ is the value of $N(s_1, s_2)$ in the graph before the i th step.

First, we define a function

$$f(d_1, d_2, t, i) = E_{t,i} \left(\sum_{s_1=1}^t \sum_{\substack{s_2=1 \\ s_2 \neq s_1}}^t [\deg s_1 = d_1, \deg s_2 = d_2] N(s_1, s_2) \right). \quad (18)$$

It is easy to see that $EX(d_1, d_2, t) = f(d_1, d_2, t, 1)$.

Recurrent equations on f are deduced as it was done in the previous sections. The sum (18) does not include the last node, so $N(s_1, s_2)$ does not change while adding a new edge. Thus, the i th step acts on f as in the case of r_2 (compare with (12)):

$$\begin{aligned} f(d_1, d_2, t, i+1) &= f(d_1-1, d_2, t, i) \frac{(d_1-1) + k(a-1)}{(a+1)(kt+i)-1} + \\ &\quad + f(d_1, d_2-1, t, i) \frac{(d_2-1) + k(a-1)}{(a+1)(kt+i)-1} + \\ &\quad + f(d_1, d_2, t, i) \left(1 - \frac{d_1 + d_2 + 2k(a-1)}{(a+1)(kt+i)-1} \right). \end{aligned} \quad (19)$$

Second, we define a function

$$g(d_1, d_2, t, i) = E_{t,i} \left([\deg(t+1) = d_2] \sum_{s=1}^t [\deg s = d_1] N(t+1, s) \right). \quad (20)$$

Obviously,

$$f(d_1, d_2, t+1, 1) = f(d_1, d_2, t, k+1) + g(d_1, d_2, t, k+1) + g(d_2, d_1, t, k+1) \quad (21)$$

and since $N(t+1, s) = 0$ before adding any edges from the node $t+1$,

$$g(d_1, d_2, t, 1) = 0. \quad (22)$$

We now consider one summand of the sum (20) and the i th step. Let the new edge link nodes $t+1$ and γ . We have three non-intersecting cases: $\gamma = s$, $\gamma = t+1$, $\gamma \notin \{s, t+1\}$. Note that

$$\begin{aligned} [\gamma = s, \deg_{t,i+1}(t+1) = d_2, \deg_{t,i+1} s = d_1] N_{t,i+1}(s) &= \\ &= [\gamma = s, \deg_{t,i}(t+1) = d_2 - 1, \deg_{t,i} s = d_1 - 1] (N_{t,i}(s) + 1), \end{aligned}$$

$$\begin{aligned} [\gamma = t + 1, \deg_{t,i+1}(t + 1) = d_2, \deg_{t,i+1} s = d_1] N_{t,i+1}(s) &= \\ &= [\gamma = t + 1, \deg_{t,i}(t + 1) = d_2 - 2, \deg_{t,i} s = d_1] N_{t,i}(s), \end{aligned}$$

$$\begin{aligned} [\gamma \notin \{s, t + 1\}, \deg_{t,i+1}(t + 1) = d_2, \deg_{t,i+1} s = d_1] N_{t,i+1}(s) &= \\ &= [\gamma \notin \{s, t + 1\}, \deg_{t,i}(t + 1) = d_2 - 1, \deg_{t,i} s = d_1] N_{t,i}(s). \end{aligned}$$

Taking the expectation and using the definition (14), we obtain

$$\begin{aligned} g(d_1, d_2, t, i + 1) &= (g(d_1 - 1, d_2 - 1, t, i) + r'_2(d_1 - 1, d_2 - 1, t, i)) \frac{(d_1 - 1) + k(a - 1)}{(a + 1)(kt + i) - 1} + \\ &\quad + g(d_1, d_2 - 2, t, i) \frac{(d_2 - 2) + (i - 1)(a - 1) + a}{(a + 1)(kt + i) - 1} + \\ &\quad + g(d_1, d_2 - 1, t, i) \left(1 - \frac{(d_1 - 1) + k(a - 1) + (d_2 - 1) + (i - 1)(a - 1) + a}{(a + 1)(kt + i) - 1} \right). \end{aligned} \quad (23)$$

Third, we derive a bound on g . Obviously, $g(d_1, d_2, t, i) = 0$ when $d_2 > 2(i - 1)$ or $d_1 \geq 2(kt + i)$. If $d_1 < 2(kt + i)$ and $d_2 \leq 2(i - 1)$, then $\frac{d_1 + d_2}{t} \cdot O(d_1^{-a-1}) = O(d_1^{-a} t^{-1}) = O(t^{-1})$ and $\frac{d_1 + d_2}{t} \cdot O(t^{-1}) = O(t^{-1})$. Remember that we have proved the bound (15) on r'_2 . It is easy to see now that

$$g(d_1, d_2, t, i + 1) = i[d_2 = i]c(d_1 - 1) \frac{(d_1 - 1) + k(a - 1)}{(a + 1)k} + O\left(\frac{1}{t}\right). \quad (24)$$

Finally, we are ready to study f . For the rest of the proof, we will assume that $d_1 \geq k$ and $d_2 \geq k$.

We denote $A = \frac{\Gamma(a+2)}{B(ka, a+1)} = (a + 1) \frac{\Gamma(ka+a+1)}{\Gamma(ka)}$, $D_1 = d_1 - k + ka$, $D_2 = d_2 - k + ka$ for brevity. By definition,

$$c(d_1) = A \frac{\Gamma(D_1)}{\Gamma(D_1 + a + 2)}.$$

Let $c_X(d_1, d_2)$ be defined recurrently as follows:

$$\begin{aligned} c_X(k, k) &= 0, \\ c_X(d_1, k) &= \frac{(D_1 - 1)(c_X(d_1 - 1, k) + c(d_1 - 1))}{D_1 + ka + a + 1}, \quad d_1 > k, \\ c_X(k, d_2) &= \frac{(D_2 - 1)(c_X(k, d_2 - 1) + c(d_2 - 1))}{D_2 + ka + a + 1}, \quad d_2 > k, \\ c_X(d_1, d_2) &= \frac{(D_1 - 1)c_X(d_1 - 1, d_2) + (D_2 - 1)c_X(d_1, d_2 - 1)}{D_1 + D_2 + a + 1}, \quad d_1, d_2 > k. \end{aligned}$$

Let

$$\begin{aligned} c_2(d_1, d_2) &= \frac{\Gamma(D_1)\Gamma(D_2)\Gamma(D_1 + D_2 + 3)}{\Gamma(D_1 + 2)\Gamma(D_2 + 2)\Gamma(D_1 + D_2 + a + 2)}, \\ c_3(d_1, d_2) &= \frac{\Gamma(D_1)\Gamma(D_2)\Gamma(D_1 + D_2 + 1)}{\Gamma(D_1 + 1)\Gamma(D_2 + 1)\Gamma(D_1 + D_2 + a + 2)}. \end{aligned}$$

Obviously, these functions are symmetric. If $d_1 > k$,

$$\begin{aligned} \frac{c_2(d_1 - 1, d_2)}{c_2(d_1, d_2)} &= \frac{(D_1 + 1)(D_1 + D_2 + a + 1)}{(D_1 - 1)(D_1 + D_2 + 2)}, \\ \frac{c_3(d_1 - 1, d_2)}{c_3(d_1, d_2)} &= \frac{D_1(D_1 + D_2 + a + 1)}{(D_1 - 1)(D_1 + D_2)}. \end{aligned}$$

Thus, for $d_1, d_2 > k$,

$$c_2(d_1, d_2) = \frac{(D_1 - 1)c_2(d_1 - 1, d_2) + (D_2 - 1)c_2(d_1, d_2 - 1)}{D_1 + D_2 + a + 1},$$

$$c_3(d_1, d_2) = \frac{(D_1 - 1)c_3(d_1 - 1, d_2) + (D_2 - 1)c_3(d_1, d_2 - 1)}{D_1 + D_2 + a + 1},$$

$$\begin{aligned} \frac{c(d_1)}{Akac_2(d_1, k)} &= \frac{\Gamma(D_1 + 2)\Gamma(ka + 2)\Gamma(D_1 + ka + a + 2)}{ka\Gamma(D_1 + a + 2)\Gamma(ka)\Gamma(D_1 + ka + 3)} = \\ &= \frac{ka + 1}{D_1 + ka + 2} \frac{\Gamma(D_1 + 2)\Gamma(D_1 + ka + a + 2)}{\Gamma(D_1 + a + 2)\Gamma(D_1 + ka + 2)}. \end{aligned}$$

Let $(\alpha)_n = \alpha(\alpha + 1) \dots (\alpha + n - 1)$ be the Pochhammer symbol. Let ${}_2F_1(\alpha, \beta; \gamma; z) = \sum_{n=0}^{\infty} \frac{(\alpha)_n(\beta)_n}{(\gamma)_n} \frac{z^n}{n!}$ for $\gamma \neq 0, -1, -2, \dots$ be the hypergeometric function. According to [1, 15.1.1], if $\gamma - \alpha - \beta > 0$ and $|z| \leq 1$, this series converges absolutely. We quote the following formula from [1, 15.1.20]:

$${}_2F_1(\alpha, \beta; \gamma; 1) = \frac{\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)}{\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)}.$$

Thus

$$\frac{\Gamma(D_1 + 2)\Gamma(D_1 + ka + a + 2)}{\Gamma(D_1 + a + 2)\Gamma(D_1 + ka + 2)} = {}_2F_1(a, ka; D_1 + ka + a + 2; 1) = \sum_{n=0}^{\infty} \frac{(a)_n(ka)_n}{(D_1 + ka + a + 2)_n n!}.$$

Since all terms of the last series are positive and the first term is 1,

$$\frac{c(d_1)}{Akac_2(d_1, k)} \geq \frac{ka + 1}{D_1 + ka + 2}.$$

Moreover,

$$\begin{aligned} \frac{\Gamma(D_1 + 2)\Gamma(D_1 + ka + a + 2)}{\Gamma(D_1 + a + 2)\Gamma(D_1 + ka + 2)} &= 1 + \sum_{n=0}^{\infty} \frac{(a)_{n+1}(ka)_{n+1}}{(D_1 + ka + a + 2)_{n+1}(n+1)!} = \\ &= 1 + \frac{a^2 k}{D_1 + ka + a + 2} \sum_{n=0}^{\infty} \frac{(a+1)_n(ka+1)_n}{(D_1 + ka + a + 3)_n(n+1)!} \leq \\ &\leq 1 + \frac{a^2 k}{D_1 + ka + a + 2} {}_2F_1(a+1, ka+1; 2ka+a+3; 1) \leq \\ &\leq 1 + \frac{a^2 k}{D_1 + ka + 1} \frac{\Gamma(ka+1)\Gamma(2ka+a+3)}{\Gamma(2ka+2)\Gamma(ka+a+2)}, \\ \frac{c(d_1)}{Akac_2(d_1, k)} &\leq \frac{ka + 1}{D_1 + ka + 2} \left(1 + \frac{kaB}{D_1 + ka + 1} \right), \quad B = a \frac{\Gamma(ka+1)\Gamma(2ka+a+3)}{\Gamma(2ka+2)\Gamma(ka+a+2)}. \end{aligned}$$

In Theorem 3, we have three assertions. The first one says that $EX(d_1, d_2, t) = c_X(d_1, d_2)t + O_{a,k}(1)$. The second one gives a bound for c_X . The third one gives an asymptotic formula for c_X . Now we shall show that our function c_X admits the bound from the second assertion. This bound is equivalent to

$$Aka \left(c_2(d_1, d_2) - \left(4 - \frac{2}{1+ka} \right) c_3(d_1, d_2) \right) \leq c_X(d_1, d_2) \leq Aka(c_2(d_1, d_2) + Bc_3(d_1, d_2)). \quad (25)$$

To prove (25), we use induction on $d_1 + d_2$. If $d_1 = d_2 = k$, the right-hand side of the inequality is obvious, and its left-hand side follows from

$$\frac{c_2(k, k)}{c_3(k, k)} = \frac{(2ka+2)(2ka+1)}{(ka+1)(ka+1)} = \frac{4ka+2}{ka+1} = 4 - \frac{2}{ka+1}.$$

If $d_1 > k$ and $d_2 > k$, all the parts of (25) satisfy the same recurrent equation, so (25) follows from the induction hypothesis. Due to symmetry, it remains to prove (25) for $d_2 = k, d_1 > k$. We have

$$\frac{ka(c_2(d_1, k) - \frac{2+4ka}{1+ka}c_3(d_1, k))}{kac_2(d_1, k)} = 1 - \frac{(D_1 + 1)(2 + 4ka)}{(D_1 + ka + 2)(D_1 + ka + 1)} = \frac{(D_1 - ka)(D_1 - ka + 1)}{(D_1 + ka + 1)(D_1 + ka + 2)}.$$

In particular, $ka(c_2(d_1, k) - \frac{2+4ka}{1+ka}c_3(d_1, k)) > 0$ for $d_1 > k$. Then,

$$\begin{aligned} \frac{c_X(d_1, k)}{Aka(c_2(d_1, k) - \frac{2+4ka}{1+ka}c_3(d_1, k))} &= \\ &= \frac{(D_1 + ka + 1)(D_1 + 1)(D_1 + ka + a + 1)}{(D_1 - ka)(D_1 - ka + 1)(D_1 - 1)} \frac{c_X(d_1, k)}{Akac_2(d_1 - 1, k)} = \\ &= \frac{(D_1 + ka + 1)(D_1 + 1)}{(D_1 - ka)(D_1 - ka + 1)} \frac{c_X(d_1 - 1, k) + c(d_1 - 1)}{Akac_2(d_1 - 1, k)} \geq \\ &\geq \frac{(D_1 + ka + 1)(D_1 + 1)}{(D_1 - ka)(D_1 - ka + 1)} \left(\frac{(D_1 - 1 - ka)(D_1 - ka)}{(D_1 + ka)(D_1 + ka + 1)} + \frac{ka + 1}{D_1 + ka + 1} \right) = \\ &= \frac{D_1 + 1}{D_1 - ka + 1} \left(\frac{D_1 - 1 - ka}{D_1 + ka} + \frac{ka + 1}{D_1 - ka} \right). \end{aligned}$$

Furthermore,

$$\begin{aligned} \frac{D_1 - 1 - ka}{D_1 + ka} + \frac{ka + 1}{D_1 - ka} - \frac{D_1 - ka + 1}{D_1 + 1} &= \frac{ka + 1}{D_1 - ka} - \frac{(ka + 1)D_1 - k^2a^2 + 2ka + 1}{(D_1 + 1)(D_1 + ka)} \geq \\ &\geq (ka + 1) \left(\frac{1}{D_1 - ka} - \frac{D_1 + ka + 1}{(D_1 + 1)(D_1 + ka)} \right) = (ka + 1)ka \frac{D_1 + ka + 2}{(D_1 + 1)(D_1 + ka)(D_1 - ka)} > 0. \end{aligned}$$

This proves the left hand-side of the inequality (25).

Now,

$$\begin{aligned} \frac{c_X(d_1, k)}{Aka(c_2(d_1, k) + Bc_3(d_1, k))} &= \\ &= \left(1 + B \frac{(D_1 + 1)(ka + 1)}{(D_1 + ka + 1)(D_1 + ka + 2)} \right)^{-1} \frac{(D_1 + 1)(D_1 + ka + a + 1)}{(D_1 - 1)(D_1 + ka + 2)} \frac{c_X(d_1, k)}{Akac_2(d_1 - 1, k)} \leq \\ &\leq \frac{(D_1 + ka + 1)(D_1 + 1)}{(D_1 + ka + 1)(D_1 + ka + 2) + B(D_1 + 1)(ka + 1)} \times \\ &\times \left(1 + B \frac{D_1(ka + 1)}{(D_1 + ka)(D_1 + ka + 1)} + \frac{ka + 1}{D_1 + ka + 1} \left(1 + \frac{kaB}{D_1 + ka} \right) \right). \end{aligned}$$

Furthermore,

$$\begin{aligned} \left(1 + B \frac{D_1(ka + 1)}{(D_1 + ka)(D_1 + ka + 1)} + \frac{ka + 1}{D_1 + ka + 1} \left(1 + \frac{kaB}{D_1 + ka} \right) \right) - \\ - \left(\frac{D_1 + ka + 2}{D_1 + 1} + B \frac{ka + 1}{D_1 + ka + 1} \right) = \left(\frac{ka + 1}{D_1 + ka + 1} - \frac{ka + 1}{D_1 + 1} \right) < 0. \end{aligned}$$

This proves the right-hand side of the inequality (25).

The third assertion of Theorem 3 (i.e., the asymptotic formula for $c_X(d_1, d_2)$) is derived from (25) in the same way as a similar formula was derived from (7).

It remains to prove the first assertion. We will use the bound $c_X(d_1, d_2) = O\left(\frac{(d_1 + d_2)^{1-a}}{d_1^2 d_2^2}\right)$.

Let

$$\tilde{f}(d_1, d_2, t, i) = f(d_1, d_2, t, i) - c_X(d_1, d_2) \left(t + \frac{i}{k} - \frac{1}{k(a + 1)} \right).$$

Let $C = C(a, k)$ be a sufficiently large constant which will be determined later. We claim that

$$\left| \tilde{f}(d_1, d_2, t, i) + (i-1) \left([d_1 = k] \frac{(D_2 - 1)c(d_2 - 1)}{(a+1)k} + [d_2 = k] \frac{(D_1 - 1)c(d_1 - 1)}{(a+1)k} \right) \right| \leqslant \leqslant C \left(1 - \frac{1}{(a+2)k(t+1)} \right)^{i-1}. \quad (26)$$

Since $r'_2(d_1, d_2, t, i) = 0$ when $d_1 + d_2 > kt + k + 2(i-1)$, (22) and (23) imply that $g(d_1, d_2, t, i) = 0$ when $d_1 + d_2 > kt + k + 2(i-1)$. Consequently, (19) and (21) imply that $f(d_1, d_2, t, i) = 0$ when $d_1 + d_2 > kt + 2k + (i-1)$.

If $d_1 + d_2 > kt + 2k + (i-1)$, then

$$\tilde{f}(d_1, d_2, t, i) = -c_X(d_1, d_2) \left(t + \frac{i}{k} - \frac{1}{k(a+1)} \right) = O \left(\frac{(d_1 + d_2)^{2-a}}{d_1^2 d_2^2} \right),$$

so (26) holds for all sufficiently large values of C .

Now assume $d_1 + d_2 \leqslant kt + 2k + (i-1)$. We will prove (26) by induction on t and, for fixed t , on i . The basis of induction $t = 1, \dots, 2 + \lfloor \frac{1}{ka} \rfloor$ and any $i = 1, \dots, k+1$ obviously holds for all sufficiently large values of C .

Now let $t \geqslant 3 + \lfloor \frac{1}{ka} \rfloor$ and let (26) hold for $t-1$. Since (26) is trivial for $d_1 = d_2 = k$ and symmetrical, we may assume that $d_1 > k$. From (21) and (24), we obtain

$$\left| \tilde{f}(d_1, d_2, t, 1) \right| = \left| \tilde{f}(d_1, d_2, t-1, k+1) + k[d_2 = k]c(d_1 - 1) \frac{D_1 - 1}{(a+1)k} + O \left(\frac{1}{t} \right) \right| \leqslant C \left(1 - \frac{1}{(a+2)kt} \right)^k + O \left(\frac{1}{t} \right).$$

The right-hand side is less than C for all sufficiently large values of C .

Finally, let $t \geqslant 3 + \lfloor \frac{1}{ka} \rfloor > 2 + \frac{1}{ka}$, $i > 1$ and let (26) hold for $i-1$. We reuse the notation $T = (a+1)(kt+i-1) - 1$ and again assume $d_1 > k$. From (19) we obtain

$$\begin{aligned} & \tilde{f}(d_1 - 1, d_2, t, i-1) \frac{D_1 - 1}{T} + [d_2 > k] \tilde{f}(d_1, d_2 - 1, t, i-1) \frac{D_2 - 1}{T} + \\ & + \tilde{f}(d_1, d_2, t, i-1) \left(1 - \frac{D_1 + D_2}{T} \right) = \tilde{f}(d_1, d_2, t, i) + c_X(d_1, d_2) \left(\frac{T}{(a+1)k} + \frac{1}{k} \right) - \\ & - \frac{c_X(d_1 - 1, d_2)(D_1 - 1) + [d_2 > k]c_X(d_1, d_2 - 1)(D_2 - 1) + c_X(d_1, d_2)(T - (D_1 + D_2))}{(a+1)k} = \\ & = \tilde{f}(d_1, d_2, t, i) + [d_2 = k]c(d_1 - 1) \frac{D_1 - 1}{(a+1)k}. \end{aligned}$$

The assumptions $d_1 + d_2 \leqslant kt + 2k + (i-1)$ and $t > 2 + \frac{1}{ka}$ imply that $1 - \frac{D_1 + D_2}{T} = 1 - \frac{d_1 + d_2 + 2k(a-1)}{(a+1)(kt+i-1)-1} \geqslant 0$. Since $(D_1 - 2)c(d_1 - 2) = [d_1 > k+1](D_1 + a)c(d_1 - 1)$,

$$\begin{aligned} & (i-2)[d_2 = k] \frac{(D_1 - 2)c(d_1 - 2)}{(a+1)k} \frac{D_1 - 1}{T} + (i-2)[d_2 = k] \frac{(D_1 - 1)c(d_1 - 1)}{(a+1)k} \left(1 - \frac{D_1 + D_2}{T} \right) = \\ & = (i-2)[d_2 = k] \frac{(D_1 - 1)c(d_1 - 1)}{(a+1)k} \left([d_1 > k+1] \frac{D_1 + a}{T} + 1 - \frac{D_1 + D_2}{T} \right) = \\ & = (i-2)[d_2 = k] \frac{(D_1 - 1)c(d_1 - 1)}{(a+1)k} + O \left(\frac{1}{t} \right). \end{aligned}$$

If $d_2 > k$, then $\frac{D_1 - 1}{T} + \frac{D_2 - 1}{T} + 1 - \frac{D_1 + D_2}{T} = 1 - \frac{2}{T} \leqslant 1 - \frac{1}{T}$. If $d_2 = k$, then $\frac{D_1 - 1}{T} + 1 - \frac{D_1 + D_2}{T} = 1 - \frac{ka+1}{T} \leqslant 1 - \frac{1}{T}$. Thus,

$$\left| \tilde{f}(d_1, d_2, t, i) + (i-1)[d_2 = k] \frac{(D_1 - 1)c(d_1 - 1)}{(a+1)k} \right| \leqslant$$

$$\leq C \left(1 - \frac{1}{(a+2)k(t+1)}\right)^{i-2} \left(1 - \frac{1}{(a+1)k(t+1)}\right) + O\left(\frac{1}{t}\right).$$

For all sufficiently large values of C , the right-hand side is less than $C \left(1 - \frac{1}{(a+2)k(t+1)}\right)^{i-1}$, so the induction on i is complete.

Theorem 3 follows from the proven bound (26).

6 Proof of Theorem 4

We use the Azuma–Hoeffding inequality.

Theorem 5. [2], [13] *Let $(X_s)_{s=0}^n$ be a martingale with $|X_{s+1} - X_s| \leq \delta$ for $s = 0, \dots, n-1$, and $x > 0$. Then*

$$P(|X_n - X_0| \geq x) \leq 2 \exp\left(-\frac{x^2}{2c^2n}\right).$$

We fix d_1, d_2, t and denote $X = X(d_1, d_2, t)$. Let G be a random graph in $H_{a,k}^{(t)}$; it has kt edges, sorted by the creation time. Let $G^{(s)}$ be a graph with s first edges. Let $X_s = E(X|G^{(s)})$, $s = 0, \dots, kt$. In this sequence $X_0 = EX$, $X_{kt} = X$. By definition of the probabilistic space, the sequence X_s is a martingale. We will estimate possible differences between adjacent elements of the sequence.

We fix any s from 0 to $kt-1$. Let v be the head of the last edge in $G^{(s+1)}$, so v is a random quantity depending on G . By definition

$$\begin{aligned} X_s &= \sum_{\gamma} \Pr(v = \gamma) E(X|G^{(s)}, v = \gamma), \\ X_{s+1} &= E(X|G^{(s)}, v = v(G^{(s+1)})), \end{aligned}$$

where the sum is over all nodes of G . Hence it is clear that

$$\min_{\gamma} E(X|G^{(s)}, v = \gamma) \leq X_s, X_{s+1} \leq \max_{\gamma} E(X|G^{(s)}, v = \gamma),$$

$$|X_s - X_{s+1}| \leq \max_{\gamma} E(X|G^{(s)}, v = \gamma) - \min_{\gamma} E(X|G^{(s)}, v = \gamma).$$

Let $\gamma_1 \in \arg \min E(X|G^{(s)}, v = \gamma)$ and $\gamma_2 \in \arg \max E(X|G^{(s)}, v = \gamma)$. It is sufficient to prove an upper bound for

$$E(X|G^{(s)}, v = \gamma_2) - E(X|G^{(s)}, v = \gamma_1).$$

We consider the sum

$$X = \sum_{s_1=1}^t \sum_{\substack{s_2=1 \\ s_2 \neq s_1}}^t [\deg s_1 = d_1, \deg s_2 = d_2] N(s_1, s_2). \quad (27)$$

Replacing the condition $v = \gamma_1$ by the condition $v = \gamma_2$ changes distributions of degrees of γ_i and distributions of $N(\gamma_i, *) = N(*, \gamma_i)$; distributions of other values of N do not change. Thus distributions of all terms in the sum (27) except those with $\{\gamma_1, \gamma_2\} \cap \{s_1, s_2\} \neq \emptyset$ are the same for $v = \gamma_1$ and $v = \gamma_2$. Let

$$X' = \sum_{s_1=1}^t \sum_{\substack{s_2=1 \\ s_2 \neq s_1 \\ \{s_1, s_2\} \cap \{\gamma_1, \gamma_2\} \neq \emptyset}}^t [\deg s_1 = d_1, \deg s_2 = d_2] N(s_1, s_2).$$

Then

$$E(X - X'|G^{(s)}, v = \gamma_1) = E(X - X'|G^{(s)}, v = \gamma_2).$$

Obviously, $X' \geq 0$. We have

$$\begin{aligned}
X' &\leq \sum_{s_1=1}^t [\deg s_1 = d_1, \deg \gamma_1 = d_2] N(s_1, \gamma_1) + \sum_{s_1=1}^t [\deg s_1 = d_1, \deg \gamma_2 = d_2] N(s_1, \gamma_2) + \\
&\quad + \sum_{s_2=1}^t [\deg \gamma_1 = d_1, \deg s_2 = d_2] N(\gamma_1, s_2) + \sum_{s_2=1}^t [\deg \gamma_2 = d_1, \deg s_2 = d_2] N(\gamma_2, s_2) \leq \\
&\quad \leq [\deg \gamma_1 = d_2] \sum_{s_1=1}^t N(s_1, \gamma_1) + [\deg \gamma_2 = d_2] \sum_{s_1=1}^t N(s_1, \gamma_2) + \\
&\quad + [\deg \gamma_1 = d_1] \sum_{s_2=1}^t N(\gamma_1, s_2) + [\deg \gamma_2 = d_1] \sum_{s_2=1}^t N(\gamma_2, s_2) = \\
&= [\deg \gamma_1 = d_2] d_2 + [\deg \gamma_2 = d_2] d_2 + [\deg \gamma_1 = d_1] d_1 + [\deg \gamma_2 = d_1] d_1 \leq 2(d_1 + d_2).
\end{aligned}$$

Thus,

$$\begin{aligned}
0 &\leq E(X'|G^{(s)}, v = \gamma_1), E(X'|G^{(s)}, v = \gamma_2) \leq 2(d_1 + d_2), \\
|E(X'|G^{(s)}, v = \gamma_1) - E(X'|G^{(s)}, v = \gamma_2)| &\leq 2(d_1 + d_2), \\
|X_s - X_{s+1}| &\leq 2(d_1 + d_2).
\end{aligned}$$

Consequently, the sequence (X_s) satisfies the condition of Theorem 5 with $n = kt$ and $\delta = 2(d_1 + d_2)$. Substituting $x = c(d_1 + d_2)\sqrt{kt}$ in Theorem 5, we obtain Theorem 4.

References

- [1] M. Abramowitz, I. A. Stegun (editors), *Handbook of mathematical functions with formulas, graphs and mathematical tables*, Dover, 1964, tenth GPO printing.
- [2] K. Azuma, *Weighted sums of certain dependent variables*, Tôhoku Math. J., 19 (1967), 357–367.
- [3] A.-L. Barabási, R. Albert, *Emergence of scaling in random networks*, Science, 286 (1999), 509–512.
- [4] B. Bollobás, *Random graphs*, Cambridge Univ. Press, Second Edition, 2001.
- [5] B. Bollobás, O. M. Riordan, *Mathematical results on scale-free random graphs*, Handbook of graphs and networks, 1–34, Wiley-VCH, Weinheim, 2003.
- [6] B. Bollobás, O. M. Riordan, J. Spencer, G. Tusnády, *The degree sequence of a scale-free random graph process*, Random Structures Algorithms, 18 (2001), N3, 279–290.
- [7] P. G. Buckley, D. Osthus, *Popularity based random graph models leading to a scale-free degree sequence*, Discrete Math., 282 (2004), 53–68.
- [8] S. N. Dorogovtsev, J. F. F. Mendes, A. N. Samukhin, *Structure of growing networks with preferential linking*, Phys. Rev. Lett. 85 (2000), 4633.
- [9] E. Drinea, M. Enachescu, M. Mitzenmacher, *Variations on random graph models for the web*, technical report, Harvard University, Department of Computer Science, 2001.
- [10] R. Durrett, *Random graph dynamics*, Cambridge, 2007.
- [11] S. Janson, T. Łuczak, A. Ruciński, *Random graphs*, Wiley, NY, 2000.
- [12] J. Jordan, *The degree sequences and spectra of scale-free random graphs*, Random Structures and Algorithms, 28, 2006.
- [13] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc., 58 (1963), 13–30.